

# **Conceptualizing and Measuring Fidelity of Implementation of Secondary Mathematics Textbooks: Results of a Three-Year Study**

Melissa D. McNaught, *University of Iowa*  
James E. Tarr, *University of Missouri — Columbia*  
Ruthmae Sears, *University of Missouri — Columbia*

*Please address all correspondence to:*

James E. Tarr  
University of Missouri  
College of Education  
121E Townsend Hall  
Columbia, MO 65211-2400  
578-882-4521  
tarrj@missouri.edu

Paper presented at the  
Annual Meeting of the American Educational Research Association

**Denver, April 30–May 4, 2010**

# **Conceptualizing and Measuring Fidelity of Implementation of Secondary Mathematics Textbooks: Results of a Three-Year Study<sup>1</sup>**

**Melissa D. McNaught**, *University of Iowa*

**James E. Tarr** and **Ruthmae Sears**, *University of Missouri*

To develop an understanding of how secondary students learn mathematics through the use of curricula, researchers must consider the way in which the materials are used, including how closely aligned classroom instruction is to the intent of the textbook author, as conveyed in the teacher guides. The National Research Council (NRC) (2004) called this alignment “implementation fidelity” and defined it as “a measure of the basic extent of use of the curricular materials” (p. 114). George, Hall, and Uchiyama (2000) suggested that implementation is a vital component when evaluating curricula. The NRC expanded on this notion, stating that evaluations of curricula require a measure of implementation in order to draw causal inferences on student achievement. More specifically, they stated,

Evaluations should present evidence that provides reliable and valid indicators of the extent, quality, and type of the implementation of the materials. At a minimum, there should be documentation of the extent of coverage of curricular materials (what some investigators referred to as “opportunity to learn”) and the extent and type of professional development provided. (p. 194)

Teachers in large part determine the implemented curriculum; they are influenced by experiences that occur within the mathematics classroom as well as by the instructional materials available to them (Ben-Peretz, 1990; Clandinin & Connelly, 1992; Clarke, Clarke, & Sullivan, 1996; Remillard, 1999; Remillard and Bryans, 2004). As such, the ways in which teachers use the same curriculum vary. Researchers need to document curriculum materials use is supported by research suggesting that curriculum implementation is an uneven process within and across schools (Grouws, 1992; Grouws & Smith, 2000; Jackson, 1992; Kilpatrick, 2003; Senk & Thompson, 2003).

This paper reports and discusses findings related to curriculum implementation in mathematics classrooms. The primary focus is on the extent and manner in which teachers use textbooks in their daily teaching. Specific attention is given to implementation differences related to two types of mathematics textbooks where the mathematical content

---

<sup>1</sup> Paper presented at the Annual Meeting of the American Education Research Association, Denver, May 2010. The authors would like to thank the following people for their assistance with all aspects of the research: Óscar Chávez, Douglas Grouws, Robert Reys, and Daniel Ross. This paper is based on research conducted as part of the Comparing Options in Secondary Mathematics: Investigating Curriculum (COSMIC) project, a research study supported by the National Science Foundation under grant number REC-0532214. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

is organized differently. We describe how multiple instruments for measuring “fidelity of implementation” (NRC, 2004) were developed, provide insights into how curriculum implementation can be measured in reliable and valid ways, report the findings of teachers’ implementation of mathematics textbooks over a three-year period, and discuss implications for researching curricular effectiveness.

The research reported in this paper is drawn from the *Comparing Options in Secondary Mathematics: Investigating Curriculum* (COSMIC) project. The goal of the COSMIC project is to examine student mathematical learning associated with secondary mathematics curriculum programs of two types – a *subject-specific* approach, where students follow a course sequence of Algebra I, Geometry, and Algebra II, and an *integrated* content approach, where students follow a course sequence of Integrated I, Integrated II, and Integrated III. To develop an understanding of the learning that occurs through the use of a particular curriculum, a large part of the project considered the way in which the curriculum materials were used in the classroom.

### **Theoretical Perspectives**

Our work is based on the assumption that the notion of fidelity of implementation has two important aspects that can be measured independently. Although fidelity is often conceived of as a measure of the extent to which teachers use written materials, there are at least two aspects of the written materials to consider: *Content Implementation* (examining *what* curriculum content is implemented) and *Presentation Implementation* (examining *how* the curriculum is implemented) (Grouws, Tarr, & McNaught, 2008). It is important to note that measures of implementation are not direct measures of the quality of teaching. That is, a low measure of implementation does not necessarily imply poor teaching.

Fidelity of implementation with regard to content is measured by determining the extent to which the content of the curriculum (textbook in this case) is used. This extent of use lies on a continuum from using the curriculum content exactly as it is written in the textbook to the other extreme of regularly skipping content or substituting content for what is in the textbook.

Presentation implementation of a curriculum refers to how textbook lessons are presented to students and the way students are expected to engage with the textbook material during the mathematics class period. This particular aspect of implementation is particularly important when studying what are often referred to as reform-oriented textbooks. For example, these textbooks often call for students to work in small group settings, to engage in discussion of ideas, to discover skills and procedures, and so on.

### **Methodology**

Evidence concerning implementation can and should be measured from several perspectives. One is that of the teacher and the other is that of the researcher. From the *teacher perspective* curriculum implementation was accessed via self-report data including

*Textbook-Use Diaries* and *Table of Contents Records*. The researcher perspective was captured through the use of *Classroom Visit Protocols*.

### ***Textbook Use Diaries (TUD)***

The TUD focused on a particular unit or chapter of the textbook and asked the teacher to provide information about lessons taught. This particular instrument asked the teacher to provide detailed information such as what materials were used to plan instruction, what supplementary materials were used in teaching the lesson, what specific examples and textbook pages were utilized during instruction, what problems were assigned as homework, and so on.

A total of 218 Textbook-Use Diaries have been collected over three years. These diaries were analyzed to describe the extent and nature of the use of the curriculum materials over 15 consecutive days of instruction. Analyses included examining how many instructional days were necessary to *cover* the specified lesson, determining what homework problems were assigned, and assessing the degree of supplementation used during the lesson.

***Number of instructional days.*** When coding diaries with respect to how many instructional days were used to cover a specified chapter or unit, a common definition of “day” must be agreed upon. In this study, a *day* was defined as a class period in a typical 7-period day with these periods ranging in length from 47 to 60 minutes. For those teachers on a block schedule, the periods ranged from 85 to 90 minutes. Conventionally, a block period is considered as two periods in a 7-period school day. Thus, the mean number of days allocated to a lesson is reported in terms of a 7-period day and data for those teachers who reported their days on a block schedule are doubled. The distributions of days allocated to each lesson were determined along with the means and standard deviations.

***Homework assignments.*** One use of the Textbook-Use Diaries was to examine the respect to which the teachers used the homework recommendations outlined by the textbook authors. The curricula for the subject-specific courses were taught from one of several textbooks with three of the most widely used published by Glencoe, Holt and McDougal-Littell. Regardless of publisher, all subject-specific textbooks recommended assignments based on three levels: Basic, Average, and Advanced. For this paper, the recommendations for the Average level were examined for these three curricula. The number of problems assigned for each section of the designated chapter for each TUD was compared to the recommendations outlined by the textbook author. The percentage of homework assignments that were fewer problems than what the authors recommended was calculated in addition to the average percent less of the problems the assignments represented.

For the integrated curriculum, the homework assignments consist of four types of problems (Modeling, Organizing, Reflecting, or Extending [MORE]) designed to complement and extend the learning from the lesson (for a full description, see McNaught, 2009). These problems are more extensive and thus result in fewer being assigned for homework. The TUDs were examined by determining the frequency of each type of problem assigned by teachers in comparison to the recommended number identified by

authors.

### ***Table of Contents Records***

The purpose of the Table of Contents (TOC) Records was to capture whether textbook lessons were taught, altered, substituted for, or skipped. Each participating teacher was provided a copy of the textbook table of contents and asked to select from four options to indicate the level of fidelity with regard to the content for each textbook section: (1) content taught primarily from textbook; (2) content taught from the textbook with some supplementation; (3) content taught primarily from an alternative source; and (4) content not taught (Figure 1).

<b><i>Unit 1 Matrix Models</i></b>		Taught primarily from <i>Core- Plus</i> textbook	Taught from <i>Core-Plus</i> textbook with some supplementation	Taught primarily from alternative(s) to <i>Core-Plus</i>	Did not teach content
<b><i>Lesson 1 Building and Using Matrix Models</i></b>					
Inv 1	<i>There's No Business Like Shoe Business</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Inv 2	<i>Analyzing Matrices</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Inv 3	<i>Combining Matrices</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 1. *Excerpt from Table of Contents Record for Integrated Textbook.*

A total of 174 Table of Contents Records were collected. Three indices were developed to capture the nature and extent of textbook use: (1) Opportunity to Learn (OTL) index; (2) Extent of Textbook Implementation (ETI) index; and (3) Textbook Content Taught (TCT) index.

In brief, the OTL index indicates whether the mathematical content contained within the textbook lessons was or was not taught. The OTL index was computed by summing the frequency of occurrence of content taught (reported across all textbook lessons on a Table of Contents Record) and then dividing by the total number of lessons included in the particular textbook. The OTL index essentially represents the percentage of the content in the textbook that students were provided an opportunity to learn.

As an example, Teacher 26, who taught from the integrated curriculum, reported 29 Investigations taught primarily from the textbook, 11 Investigations taught from the textbook with some supplementation, 9 taught primarily from an alternative source, and

28 not taught out of a total of 77 Investigations. Thus, the OTL index is calculated as follows:

$$OTL = \frac{29 + 11 + 9}{77} * 100 = 63.64$$

It should be noted that when interpreting OTL indices, an index of 63 does not imply that the coverage corresponds to the *first 63%* of the textbook. Omitted lessons occur at many places as teachers progress through the textbook (Grouws, Tarr, & McNaught, 2008).

Next, the ETI index was determined by weighting each of the first three options on the Table of Contents Record. The largest weight was given when the first option was identified for a lesson. That is, when lesson content was taught primarily from the textbook, it was assigned a weight of 1. Content not taught was given a weight of 0. The two options in between, content taught with supplementation and content taught primarily from an alternative source were assigned weights of 2/3 and 1/3, respectively. The index was then calculated by summing the weights across textbook lessons and then dividing by the number of lessons contained in the particular textbook. The quotient was then multiplied by 100 giving the ETI index a scale ranging from 0 to 100. An index of 100 would represent that every lesson contained in the textbook was taught directly from the textbook and done so without supplementation or use of alternate sources. An index of 0 would indicate that no lessons from the textbook were taught.

Finally, the TCT index differs from the ETI index by considering only those lessons where content was taught in some manner, thereby ignoring content students were not given the opportunity to learn. The lessons were weighted in the same manner as in the ETI, but the index was calculated by dividing by the number of lessons reported as being taught in any manner and again multiplied by 100. The index is reported on a scale ranging from 0 to 100. An index less than 100 would indicate that some lessons taught were taught with supplementation or the use of alternative sources. Thus indices less than 100 indicate the extent to which textbook lessons taught were supplemented or replaced. Ultimately, this index reports the extent to which teachers, *when teaching textbook content*, followed their textbook, supplemented their textbook lessons, or used altogether alternative curricular materials.

### ***Classroom Visit Protocols***

***Classroom Visit Protocols*** (Tarr, McNaught, & Sutter, 2006) were developed after studying the recommendations offered in the teacher's edition of the textbooks and interviewing authors to assemble a profile of what the curriculum developers expected of teachers in implementing their curricular materials. These protocols are curriculum specific in order to guide the classroom observer to note important content aspects of the curriculum attended to in the classroom such as exercises or examples worked on or discussed.

Each of the teachers in the study was observed at least three times during the school year for a total of 326 observations completed during the first two years of data collection.

During the lesson, the classroom visitor recorded detailed notes and then immediately after the lesson completed a protocol upon which s/he had received previous training. Upon completion of a class visit, the observer indicated the overall level of fidelity of implementation with regard to textbook content and recommended teaching process using Likert rating scales. High Content Fidelity consisted of the content being enacted as written in the textbook and the textbook used as the primary source of content for the lesson. High Presentation Fidelity indicated the enacted curriculum being consistent with the expectations of the authors as expressed in the author interviews and the author philosophy as reflected in the notes and suggestions to teachers in the Teacher's Edition of the textbook. Each rating was independent of the other allowing for the possibility of a high fidelity rating on one aspect of implementation and a low fidelity rating on the other for a given lesson. Coding reliability was studied prior to and during the study and showed high inter-rater reliability coefficients (Grouws, Tarr, & McNaught, 2008).

The instruments together (TUD and TOC) provide both the teacher and research perspectives to yield a robust picture of curriculum implementation (1) over an academic year, (2) within a specific textbook chapter, and (3) within individual lessons taught.

## **Results**

### ***Textbook Use Diaries – Instructional Days***

The first analysis of the diaries examined how many days each teacher spent on a specified chapter in the subject-specific textbooks or on a multi-day lesson in the integrated textbook as compared to the recommendations outlined by the textbook authors. The mean number of instructional days is reported in terms of a 7-period day; for those teachers who reported their days on a block schedule, the number of instructional days has been doubled.

For each chapter in each textbook, authors recommend the number of days that would be considered optimal to teach the content contained within the chapter. The average distribution of days teachers allocated for the subject-specific textbooks is shown in Table 1. As the data reveal, on average teachers spent fewer days on the chapter than the textbook authors recommended with the exception of the teachers using the McDougal-Littell Algebra 1 textbook. These teachers on average followed the recommendations more closely.

When examining teachers' use of the integrated textbooks, the situation was reversed with a tendency for teachers to spend more than the recommended number of days. Recall the format of the integrated textbook utilizes multi-day lessons and for each lesson, textbook authors recommend the optimal amount of days to allocate to the content. Table 2 displays the distribution of days allocated for the integrated textbooks. For three of the five identified lessons, teachers on average spent more days than recommended.

Table 1  
*Distribution of Days Teachers Allocated to Chapters in Subject-Specific Textbooks*

Text	Course	Chapter	# of TUDs	Number of Days Teachers Allocated				Recommended # of Days
				Min Days	Max Days	Mean Days	SD	
Glencoe	Algebra 1	3	9	4	15	12.00	3.64	16
Glencoe	Algebra 1	5	10	6	15	12.00	3.02	14
McDougal-Littell	Algebra 1	4	6	15	19	15.67	1.63	15
McDougal-Littell	Algebra 1	5	6	7	12	13.50	4.68	13
Holt	Geometry	3	8	8	16	11.50	2.98	13
McDougal-Littell	Geometry	3	24	0	16	10.38	4.20	12
Glencoe	Algebra 2	3	17	4	12	7.65	3.06	10
McDougal-Littell	Algebra 2	3	8	3	13.5	6.88	3.73	10

Table 2  
*Distribution of Days Teachers Allocated to Lessons in Integrated Textbooks*

Text	Course	Unit & Lesson	# of TUDs	Number of Days Teachers Allocated				Recommended # of Days
				Min	Max	Mean	SD	
Core-Plus	Course 1	U3L2	21	1	15	8.10	4.02	12
Core-Plus	Course 1	U6L2	20	1	20	10.95	6.00	5
Core-Plus	Course 2	U2L2	27	3	30	14.04	5.75	12
Core-Plus	Course 3	U3L4	11	0	12	6.59	4.24	7
Core-Plus	Course 4	U6L1	11	3	26	12.55	6.61	7

### ***Textbook Use Diaries –Homework Assignments***

The examination of the number of instructional days provides a general overview of the time spent on the particular content of a chapter or lesson but more specific information is revealed from an analysis of the teacher use of the homework components of textbooks.



Table 3 displays the degree to which teachers using subject-specific textbooks follow author recommendations. As these data reveal, most teachers assigned far fewer problems than the number recommended in the teacher's editions. The average deficiency column indicates what percentage of the teachers assign less than what the textbooks authors consider optimal. The percentage of deficiency was somewhat similar within the results associated with the Algebra 1 textbooks, however, increased as the course level increased.

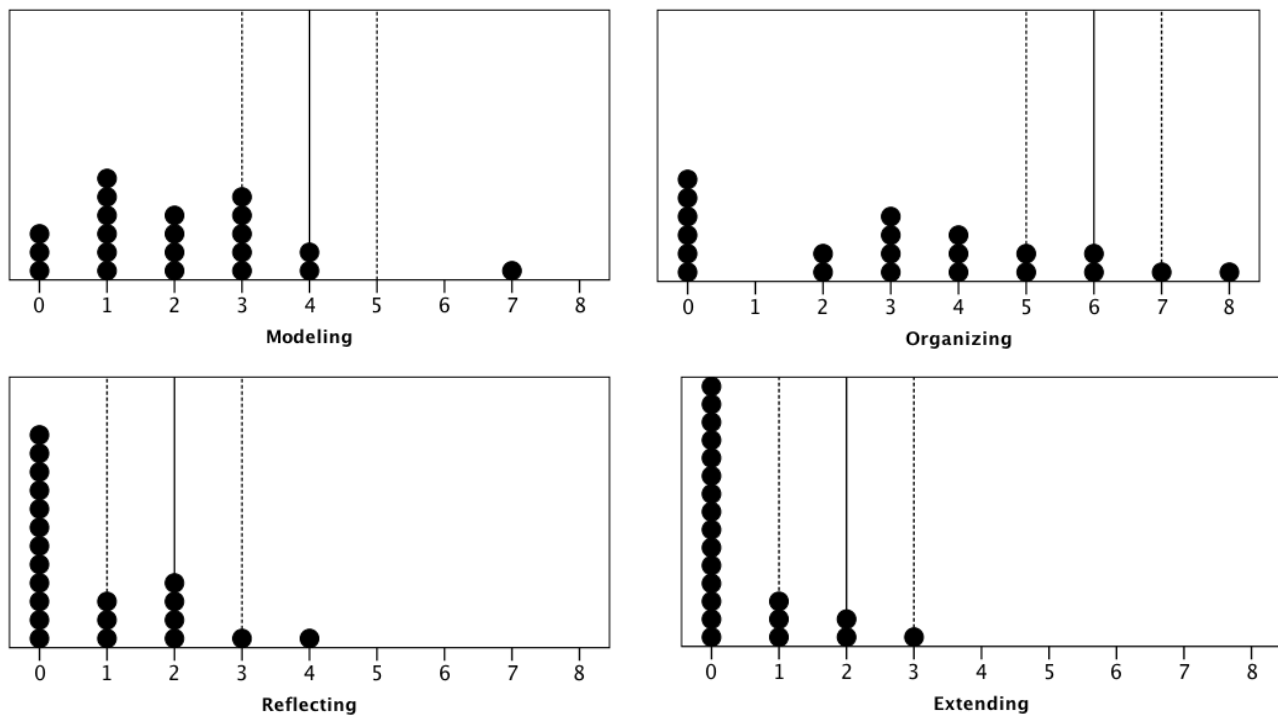
Table 3

*Teacher Use of the Homework Components of Subject- Specific Textbooks*

Text	Course	Chapter	# of Sections in Chapter	# of TUD	Total Entries	% of Lessons Less than Recommended	Average Deficiency %
Glencoe	Algebra 1	3	9	9	81	85%	51%
Glencoe	Algebra 1	5	7	10	70	83%	40%
McDougal-Littell	Algebra 1	4	8	6	48	100%	47%
McDougal-Littell	Algebra 1	5	7	6	42	100%	44%
Holt	Geometry	3	6	8	48	98%	79%
McDougal-Littell	Geometry	3	7	24	168	98%	66%
Glencoe	Algebra 2	3	5	17	85	100%	89%
McDougal-Littell	Algebra 2	3	6	8	48	100%	81%

Similar patterns were observed for teachers of the integrated curriculum, however, because of the extensive nature of the problems associated with this type of curriculum, fewer problems are generally assigned. However, authors do recommend an optimal number of problems of each type to be assigned for homework. Although the authors recommend the MORE problems as a means for assessing the level of understanding of each individual student, the results of this study indicate that teachers did not closely abide by this recommendation. Figures 2, 3 and 4 display the use of the MORE problems by teachers during each year disaggregated by MORE problem type for a common lesson. The dark line represents the author-recommended number of each particular problem type to be assigned. The dotted lines represent an interval of within  $\pm 1$  about the author recommended number.

During year 1, most teachers assigned far fewer problems than the number recommended by the textbook authors. As displayed in Figure 2, 18 of 21 teachers assigned fewer Modeling problems and 17 teachers assigned fewer Organizing problems than were advised. Among the four types of problems, the Reflecting and Extending tasks were the least assigned problems with 12 teachers not assigning any Reflecting problems and 15 teachers not assigning any Extending problems.



*Figure 2: Frequency of the Number of MORE Problems Assigned by Teachers During a Common Lesson (U3L2), Year 1.*

This trend is even more pronounced during year 2 in that 22 teachers assigned fewer Modeling problems than recommended and 25 teachers assigned fewer Organizing problems (Figure 3). Again, the Reflecting and Extending problems were rarely utilized with only 2 teachers reporting any use of the Reflecting problems and only 5 teachers reporting use of Extending problems. Moreover, when considering the  $\pm 1$  interval across the two years, teachers consistently fall on the lower side of the interval. By a ratio of 4 to 1, teachers were more likely to assign one fewer problem than one more than the number of problems recommended.

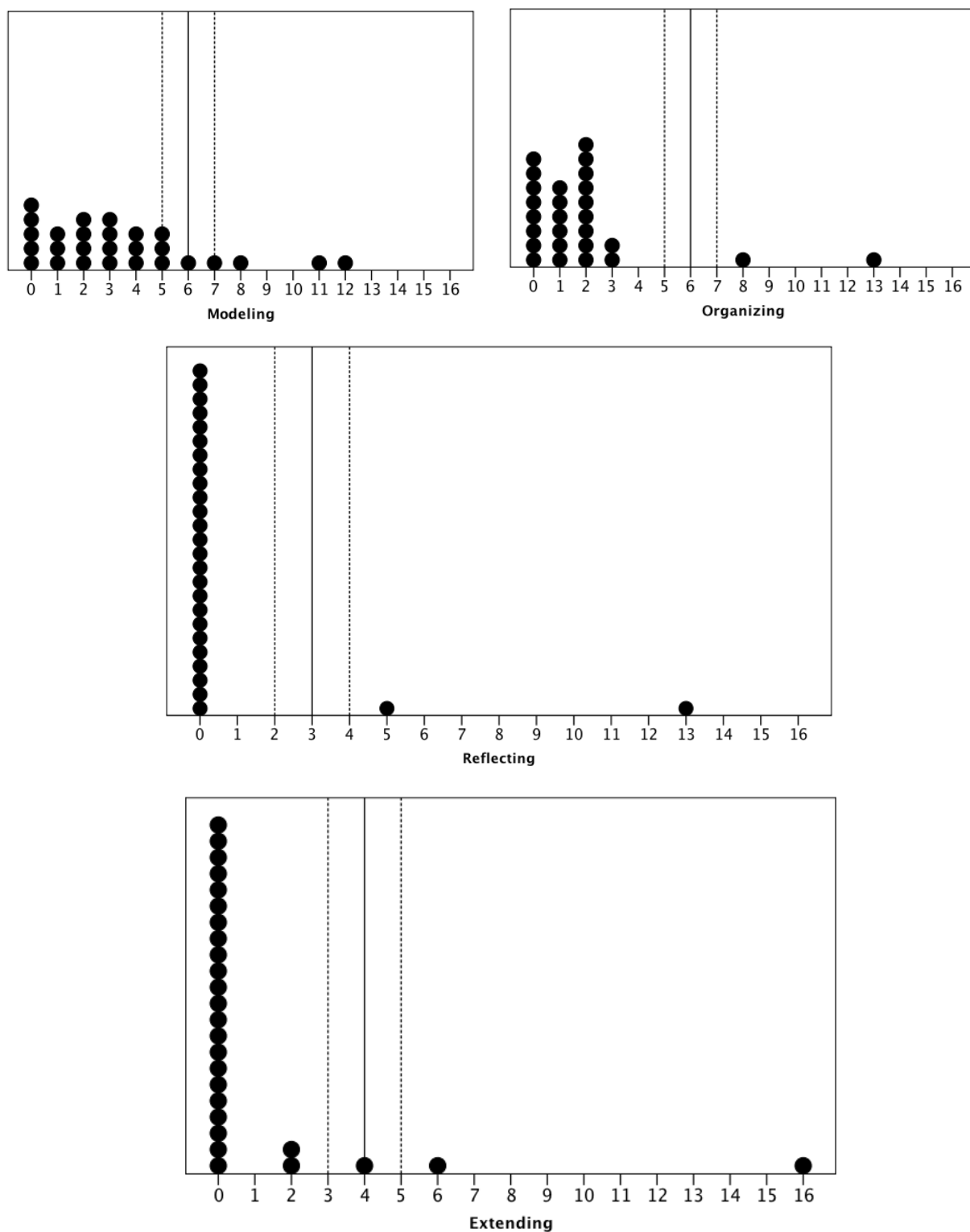
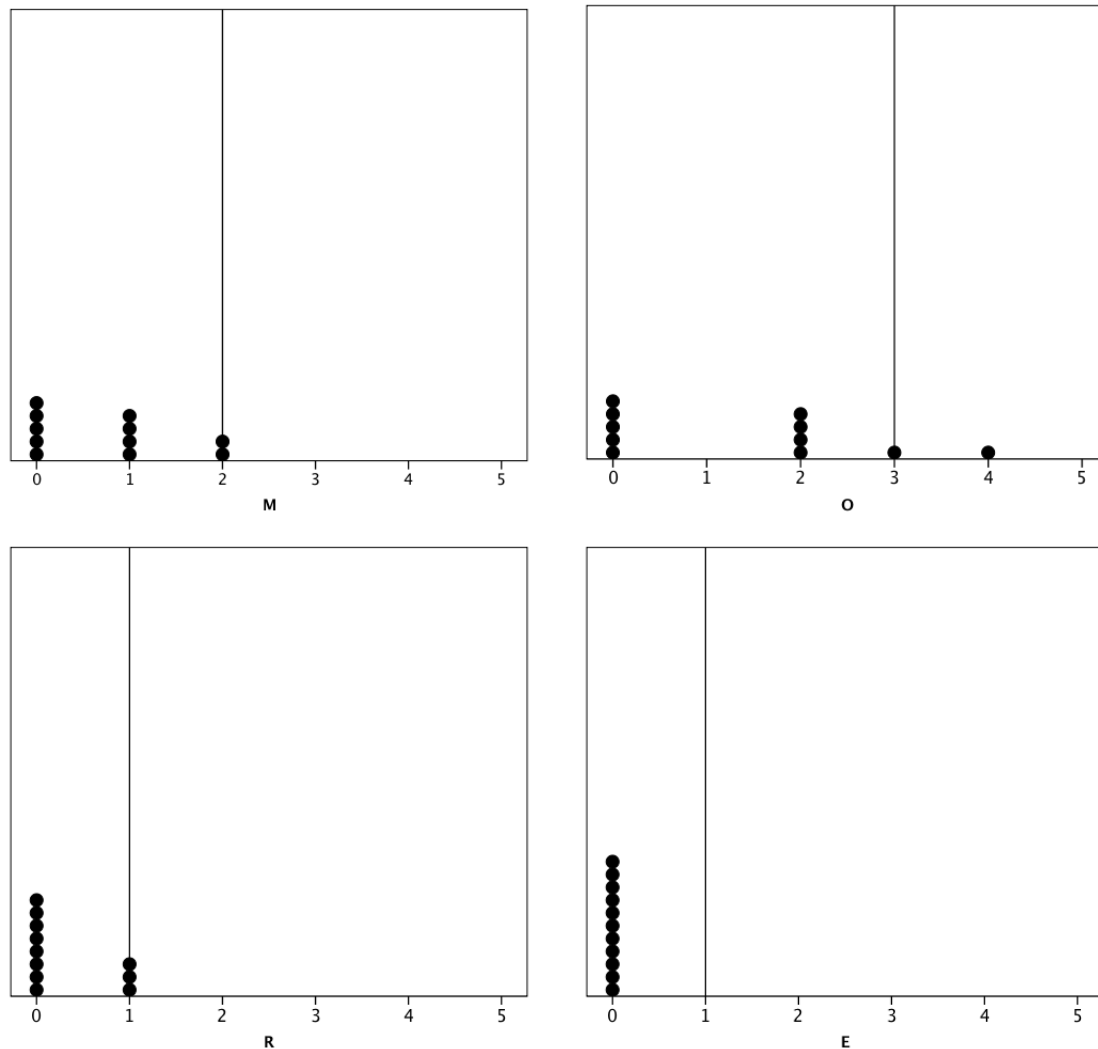


Figure 3: Frequency of the Number of MORE Problems Assigned by Teachers During a Common Lesson, Year 2.

As shown in Figure 4, the trend continued into the third course of the integrated series. Again, most teachers assigned far fewer problems than the number recommended in the teacher's editions. Specifically, 9 of 11 teachers assigned fewer Modeling problems and Organizing problems than were advised. Moreover, 8 of 11 teachers assigned no Reflecting problems whereas all eleven teachers assigned no Extending problems.

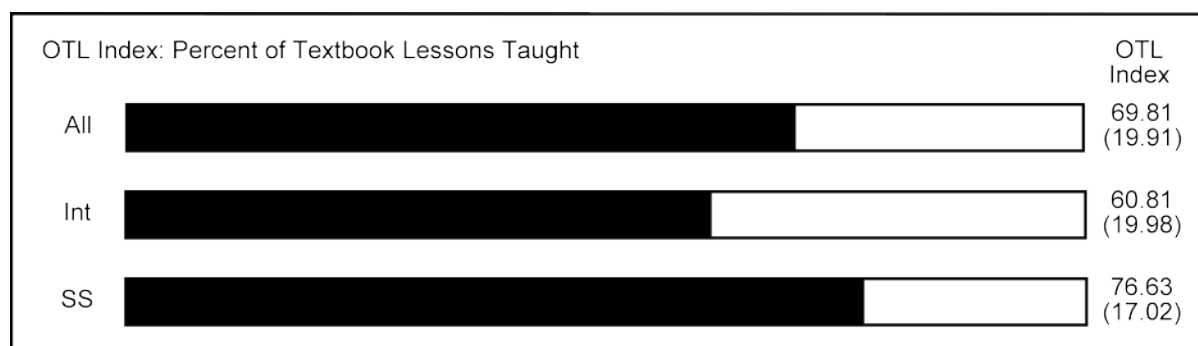


*Figure 4: Frequency of the Number of MORE Problems Assigned by Teachers During a Common Lesson (C3U3L4) Year 3.*

### ***Table of Contents Records***

Using the Table of Content Records, an overall mean opportunity to learn (OTL) index across the 174 records was computed. The mean OTL index was 69.81 with standard deviation of 19.91 as represented in the first graph of Figure 5. This indicates the percent of

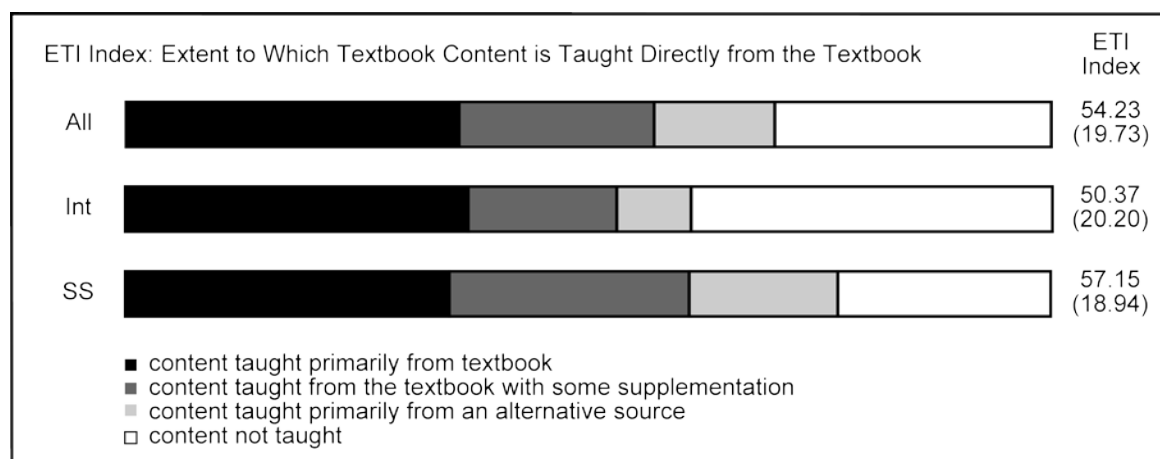
content taught *as defined by textbook lessons* that students were provided the opportunity to learn. The next two graphs display the mean percent of the content in the integrated textbook covered and the mean percent of the content in the subject-specific textbook covered, respectively. These data indicate that less than 75% of the content embodied in the textbook was covered regardless of the type of textbook in use. However, there is substantial variation among teachers in terms of the OTL index.



*Figure 5. Summary of Opportunity to Learn (OTL) indices.*

Note: Heavy shading indicates the sum of the frequency of the first three options reported on TOC: (1) *content taught primarily from textbook*; (2) *content taught from the textbook with some supplementation*; (3) *content taught primarily from an alternative source*. White indicates content not taught.

Figure 6 below displays the overall *Extent of Textbook Implementation* (ETI) index and the ETI indices disaggregated by textbook type. As the data show, across teachers slightly more than one-third (35%) of the content was taught primarily from the textbook, approximately one-fifth (21%) of the content was taught with some supplementation, a small portion (12%) was taught from alternative resources and 32% of the content was not taught. These data also indicate that the variance is similar regardless of the type of textbook the teacher was using.



*Figure 6. Summary of the Extent of Textbook Implementation (ETI) indices.*

Figure 7 displays the *Textbook Content Taught* indices. The TCT index restricted the focus to just the textbook content that was taught. This index differs from the ETI index where the unit of analysis was the entire textbook. The results show that for the content taught across all teachers, 51% of that content is taught directly from the textbook, 31% of that content was taught with some supplementation, and 18% of content taught from alternative resources yielding an overall TCT index of 77.77. As shown visually in the second and third bars of the figure, integrated content was taught more frequently directly from textbook (59%) as compared to subject-specific content (46%). Note that the difference was smaller with regard to content supplementation with teachers using the subject-specific curriculum materials supplementing 33% of the content and teachers using the integrated curriculum materials supplementing a little less (28%). In general, very little textbook content was taught from alternate sources (18% across textbook types, 13% integrated, 21% subject-specific). Interestingly the TCT variance is slightly greater for teachers using the subject-specific curriculum materials than it is for teachers using the integrated curriculum materials suggesting that subject-specific teachers feel less constrained when teaching the textbook content.

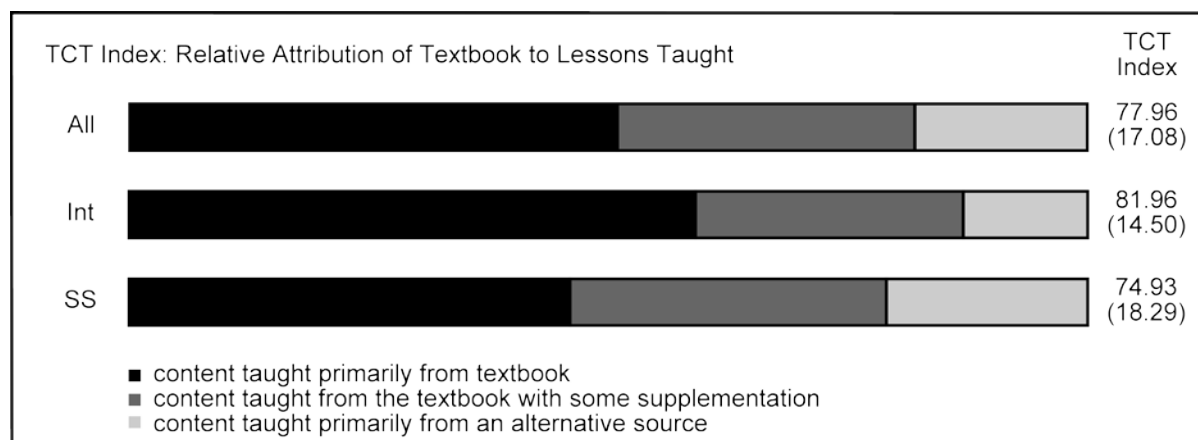


Figure 7. Summary of Textbook Content Taught (TCT) indices.

### ***Classroom Visit Protocols***

During classroom observations, trained observers recorded judgments regarding the degree to which the textbook influenced the content taught (Content Fidelity) and the manner of presentation of the mathematics lessons (Presentation Fidelity). Analyses of Content Fidelity ratings showed that the content of lessons were primarily attributable to the textbook (see Table 4). Across all teachers, the mean Content Fidelity rating was 3.67 (on a 5 point scale) across 326 lessons. Moreover, these data indicate that Content Fidelity ratings were relatively high regardless of whether the teacher was teaching from an integrated or a subject-specific textbook and the means were not significantly different across textbook types.

With regard to Presentation Fidelity, the overall ratings are lower and significantly different ( $p < .001$ ) from the Content Fidelity ratings. The mean overall Presentation Fidelity rating was 3.11 across the 109 teachers and 326 lessons suggesting that the

*manner* in which the lessons were taught was less consistent with the authors' expectations than was the *content* of lessons taught. Furthermore, the presentation rating for the integrated lessons was significantly lower than that of the subject-specific ( $p < .05$ ).

Table 4  
*Teacher Content and Presentation Fidelity*

	N	Content Fidelity		Presentation Fidelity	
		Mean	Standard Dev	Mean	Standard Dev
All	109	3.67	1.02	3.11	0.96
Integrated	48	3.88	1.06	2.91	1.11
Subject-Specific	61	3.50	0.96	3.28	0.80

### ***Relationship of Content Fidelity and Presentation Fidelity***

The correlation between the content ratings and the presentation ratings was 0.50. Thus, there appears to be a moderate relationship between the two dimensions. In other words, those teachers who tend to have higher content ratings also tend to have higher presentation ratings. However, this correlation is not so high that it suggests Content and Presentation Fidelity should be examined as a single dimension of fidelity.

### **Significance of the Study**

Plausible reasons for teachers' variation of the implementation of textbook content may be attributable to the requirements of mathematics testing outlined in No Child Left Behind. In her study, Bowzer (2008) reported the influence of state testing programs on the composition of district curriculum guides in order to ensure students' opportunity to learn key mathematical content. If district guides are not available, teachers may decide to focus on certain content because they perceive it to be the content mostly likely to be included on the test (Stecher & Barron, 2001). Furthermore, teachers may give preference to the content that they perceive students will need for college.

Given that during the course of three successive school years as much as one-third of the content includes the use of supplementary materials, any measures of the impact of these particular curricula on student learning should be re-examined, if this supplementation was not accounted for during data analysis. Measures of opportunity to learn in the past have been simple dichotomies, but these do not provide an adequate picture of the degree to which the textbook is used to teach content (McNaught, 2009). Using the three indices computed in this study, OTL, ETI, and TCT, provides a much more robust documentation of textbook use. While the OTL index provides information on whether students are learning a topic, the ETI and TCT indices provide a more refined measure of the extent to which the textbook is used in teaching mathematics content. The documentation of "fidelity of implementation" is necessary in order to make valid interpretations of student learning in relation to a specific textbook curriculum (NRC, 2004) and yet few studies of curricular effectiveness take appropriate account of curriculum implementation (O'Donnell, 2008). Our study provides one model for how textbook implementation can be measured empirically in valid and reliable ways.

Researchers have established that teachers rely heavily on curricular materials when making decisions regarding what content to teach. Furthermore, school districts can select from a large number of textbooks that differ in substantive ways from one another. Studies of how mathematics textbooks are used are needed in order for school personnel, curriculum developers, and policy makers to make informed decisions regarding curriculum. Moreover, student achievement in secondary school mathematics is significantly influenced by the manner in which a textbook is used, the mathematics content students are given the opportunity to learn, and the manner in which the mathematical learning is facilitated; however, our knowledge of the effects of different approaches to mathematical content in textbooks is limited. Furthermore, it is impossible to judge effectiveness of textbooks and draw inferences regarding student learning without knowledge about how those textbooks are being used (National Research Council, 2004).

In light of new curricular materials being developed and used in classrooms, research results garnered from examining the degree to which the textbook is being implemented and the quality of its delivery have the potential to inform curriculum developers of ways to improve their programs. Our research suggests some methodological approaches to fulfilling this need and shows that such studies can provide informative information and demonstrates that important differences exist.

## References

- Ben-Peretz, M. (1990). *The teacher-curriculum encounter: Freeing teachers from the tyranny of texts*. Albany, NY: State University of New York Press.
- Bowzer, A. (2008). *Identity and curricular construction: A study of teacher interaction with mathematics curricula of two types*. Unpublished doctoral dissertation, University of Missouri, Columbia.
- Clanindin, D.J., & Connelly, F.M. (1992). Teacher as curriculum maker. In P. Jackson (Ed.), *The handbook of research on curriculum* (pp. 363-396). New York: Macmillan.
- Clarke, B., Clarke, D., & Sullivan, P. (1996). The mathematics teacher and curriculum development. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 1207-1233). Dordrecht, The Netherlands: Kluwer.
- George, A. A., Hall, G. E., & Uchiyama, K. (2000). Extent of implementation of a *Standards-based approach to teaching mathematics and student outcomes*. *Journal of Classroom Interaction*. 35(1). 8-25.
- Grouws, D. A. (Ed.). (1992). *Handbook of research in mathematics teaching and learning*. New York: Macmillan.
- Grouws, D. A. & Smith, M. (2000). NAEP findings on the preparation and practices of mathematics teachers. In E.A Silver & P.A. Kenney (Eds.), *Results from the Seventh*



- Mathematics Assessment of the National Assessment of Educational Progress* (pp. 107-139). Reston, VA: National Council of Teachers of Mathematics.
- Grouws, D.A., Tarr, J.E., & McNaught, M. D. (2008, November). *Curriculum implementation of mathematics textbooks of different types in secondary school classrooms*. Proceedings of the annual meeting of The Australian Association of Research in Education (AARE), Brisbane, Australia.
- Jackson, P. W. (Ed.). (1992). *Handbook of research on curriculum: A project of the American Educational Research Association*. New York: Macmillan.
- Kilpatrick, J. (2003). What works? In S. L. Senk & D. R. Thompson (Eds.), *Standards-based school mathematics curricula. What are they? What do students learn?* (pp. 471-493). Mahwah, NJ: Lawrence Erlbaum.
- McNaught, M.D. (2009). *Implementation of integrated mathematics textbooks in secondary school classrooms*. Unpublished doctoral dissertation, University of Missouri, Columbia.
- National Research Council. (2004). *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics evaluations*. Washington, DC: National Academies Press.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 intervention research. *Review of Educational Research*, 78(1), 33-84.
- Remillard, J. T. (1999). Curriculum materials in mathematics education reform: A framework for examining teachers; curriculum development. *Curriculum Inquiry*, 29(3), 315-342.
- Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education*, 35, 352-388.
- Senk, S. L. & Thompson, D. R. (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Tarr, J. E., McNaught, M. D., & Sutter, A. D. (2006). *Core-Plus classroom visitation protocol*. Columbia, MO: University of Missouri.
- Tarr, J. E., McNaught, M. D., & Sutter, A. D. (2006). *Subject-specific classroom visitation protocol*. Columbia, MO: University of Missouri.
- Tarr, J. E., McNaught, M. D., & Sutter, A. D. (2006). *User's guide: Core-Plus classroom visitation protocol*. Columbia, MO: University of Missouri.
- Tarr, J. E., McNaught, M. D., & Sutter, A. D. (2006). *User's guide: Subject-specific classroom visitation protocol*. Columbia, MO: University of Missouri.